

# The Neurophysiology of Implicit Alcohol Associations in Recently Abstinent Patients With Alcohol Use Disorder: An Event-Related Potential Study Considering Gender Effects

Raphaela Martina Tschuempferlin , Hallie Margareta Batschelet , Franz Moggi ,  
Thomas Koenig , Susanne Roesner, Anne Keller, Philippe Pfeifer, Leila Maria Soravia ,  
and Maria Stein 

**Background:** Neuroscientific models of alcohol use disorders (AUDs) postulate an imbalance between automatic, implicit, and controlled (conscious) processes. Implicit associations towards alcohol indicate the automatically attributed appeal of alcohol-related stimuli. First, behavioral studies indicate that negative alcohol associations are less pronounced in patients compared to controls, but potential neurophysiological differences remain unexplored. This study investigates neurophysiological correlates of implicit alcohol associations in recently abstinent patients with AUD for the first time, including possible gender effects.

**Methods:** A total of 62 patients (40 males and 22 females) and 21 controls performed an alcohol valence Implicit Association Test, combining alcohol-related pictures with positive (incongruent condition) or negative (congruent condition) words, while brain activity was recorded using 64-channel electroencephalography. Event-related potentials (ERPs) for alcohol-negative and alcohol-positive trials were computed. Microstate analyses investigated the effects of group (patients, controls) and condition (incongruent, congruent); furthermore, possible gender effects in patients were analyzed. Significant effects were localized with standardized low-resolution brain electromagnetic topography analysis.

**Results:** Although no behavioral group differences were found, ERPs of patients and controls were characterized by distinct microstates from 320 ms onwards. ERPs between conditions differed only in patients with higher signal strength during incongruent trials. Around 600 ms, controls displayed higher signal strength than patients. A gender effect mirrored this pattern with enhanced signal strength in females as opposed to male patients. Around 690 ms, a group-by-valence interaction indicated enhanced signal strength in congruent compared to incongruent trials, which was more pronounced in controls.

**Conclusions:** For patients with AUD, the pattern, timing, and source localization of effects suggest greater effort regarding semantic and self-relevant integration around 400 ms during incongruent trials and attenuated emotional processing during the late positive potential timeframe. Interestingly, this emotional attenuation seemed reduced in female patients, thus corroborating the importance of gender-sensitive research and potential treatment of AUD.

**Key Words:** Alcohol Use Disorder, Implicit Association Test, Event-Related Potentials, Microstates, Gender.

From the Translational Research Center, (RMT, HMB, FM, TK, PP, LMS, MS), University Hospital of Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland; Clinic Suedhang., Kirchliedach, Switzerland; Graduate School for Health Sciences, (RMT, HMB), University of Bern, Bern, Switzerland; Addiction Treatment Center, (SR, AK), Forel Clinic, Ellikon an der Thur, Switzerland; and Department of Clinical Psychology and Psychotherapy, (MS), Institute of Psychology, University of Bern, Bern, Switzerland.

Received for publication February 16, 2020; accepted August 17, 2020.

Reprint requests: Raphaela M. Tschuempferlin, PhD, Translational Research Center, University Hospital of Psychiatry and Psychotherapy, Bolligenstrasse 111, CH-3000 Bern 60, Switzerland; Tel.: +41-31-9328369; Fax: +41-31-9309961; E-mail: raphaela.tschuempferlin@upd.unibe.ch

© 2020 The Authors. Alcoholism: Clinical & Experimental Research published by Wiley Periodicals LLC on behalf of Research Society on Alcoholism.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

DOI: 10.1111/acer.14444

*Alcohol Clin Exp Res*, Vol 44, No 10, 2020; pp 2031–2044

CURRENT NEUROSCIENTIFIC models of substance use disorders postulate an imbalance between automatic, implicit, and control processes (e.g., Schacht et al., 2013; Volkow and Baler, 2014). Here, implicit associations toward alcohol, as measured with the Implicit Association Test (IAT; Greenwald et al., 1998, 2003), could be an indicator for this automatically attributed appeal of alcohol-related stimuli. The IAT measures the relative strength of associations between concepts (indicated by the reaction time-based IAT effect) and is based on the rationale that people are slower in combining incongruent versus congruent concepts.

Preclinical behavioral research using bipolar alcohol valence IATs indicated strong negative alcohol associations, which were less negative in heavy than in light drinkers (Wiers et al., 2005; Wiers et al., 2002). However, unipolar IATs yielded positive as well as negative implicit alcohol associations, but positive associations correlated stronger

with alcohol use (e.g., Ames et al., 2014; Houben and Wiers, 2008; Jajodia and Earleywine, 2003; Wiers et al., 2002). In clinical samples, unipolar IATs yielded positive and negative associations, whereby only negative associations were weaker in patients than in controls (Dickson et al., 2013). A bipolar IAT revealed a negative bias in patients with alcohol use disorders (AUD; De Houwer et al., 2004). Based on these first results, we (i) define the combination of alcohol-related pictures with negative word as the congruent condition, whereas in the incongruent condition, positive words are paired with alcohol-related pictures; and (ii) we expect less negative alcohol associations in patients than controls.

While behavioral research has begun to explore alterations in the reaction time-based IAT effects in patients with AUD, the neurophysiological basis of the alcohol-related IAT effect remains unknown. Previous neurophysiological studies on the IAT mostly focused on various aspects of social cognition (e.g., Knutson et al., 2007; Schiller et al., 2016; Williams and Themanson, 2011) or self-evaluative processes (e.g., Egenolf et al., 2013; Fleischhauer et al., 2014; Xiao et al., 2015). Only one study examined event-related potentials (ERPs) of an addiction-related IAT in a sample of patients with Internet addiction (Chen et al., 2018). Along with reports on early components (e.g., P1; Fleischhauer et al., 2014), those studies have repeatedly indicated that the amplitude (Chen et al., 2018; Coates and Campbell, 2010; Xiao et al., 2015) and timing (Schiller et al., 2016) of the N2 component differ between incongruent and congruent trials, which was attributed to a larger prereponse conflict in incongruent trials.

Regarding later components, reports included higher amplitudes (O'Toole and Barnes-Holmes, 2009; Williams and Themanson, 2011) during the N400 timeframe, which typically varies with semantic congruency and might reflect sensibility to semantic violations in incongruent conditions.

Further, studies indicated that the late positive potential (LPP), a central positivity, linked to sustained and more elaborative processing of salient emotional stimuli (Cuthbert et al., 2000; Dillon et al., 2006), varies with IAT conditions (Hurtado et al., 2009; O'Toole and Barnes-Holmes, 2009; Williams and Themanson, 2011). These studies suggested that enhanced LPP in congruent trials reflects stronger and/or longer emotional and evaluative processing. Finally, Schiller and colleagues (2016) reported a longer microstate representing cognitive control in the selection of motor responses.

Source analysis of ERP data indicated that brain regions involved in IAT effects include the lingual gyrus as a generator of early effects (Schiller et al., 2016) and midcingulate, posterior parietal, and frontal regions as generators of late effects (Egenolf et al., 2013; Schiller et al., 2016). fMRI studies on IAT effects in social cognition revealed higher activation in ventrolateral and dorsolateral prefrontal cortex and anterior cingulate (Chee et al., 2000; Knutson et al., 2006; Luo et al., 2006) during incongruent compared to congruent trials, highlighting the role of inhibitory processes during the processing of these trials. Congruent, in contrast to

incongruent, trials activated brain areas such as orbitofrontal, medial temporal, and insular regions, but superior and medial frontal activation was reported as well (Knutson et al., 2007; Knutson et al., 2006).

Two fMRI studies investigated brain activation during substance-related IATs with preclinical samples: Ames and colleagues (2013) found significantly greater bilateral activity in the caudate and putamen in marijuana users during performance of congruent trials relative to controls. During incongruent trials, controls showed enhanced activity in the right inferior frontal gyrus region compared to users. Performing a unipolar alcohol valence IAT, light drinkers exhibited higher activation in the left orbitofrontal cortex compared to heavy drinkers (Ames et al., 2014). In heavy drinkers, the left putamen and insula were enhanced in alcohol-positive compared to alcohol-neutral mappings.

As the neurophysiological basis of implicit alcohol-related associations in a clinical sample with AUD is still unknown, the first aim of the present study is to investigate ERPs recorded with multichannel electroencephalography (EEG) during an alcohol valence IAT. Microstate analysis segments the continuous ERP signal into a series of quasistable topographies, each representing a specific step of mental processing by distinct underlying networks (Koenig et al., 2013; Lehmann and Skrandies, 1980). Such analyses allow to analyze ERPs in a data-driven, reference-independent way, without a priori assumptions regarding generators or preferable electrode positions (Koenig et al., 2013), and have successfully been applied to ERPs collected during an IAT (Schiller et al., 2016). Applying these analyses to compare the neurophysiological activation of patients with AUD to healthy controls will broaden our understanding of neuronal underpinnings of implicit alcohol associations in AUD.

The second aim of this study is to examine gender effects. Recent research and societal developments highlight the importance of gender-sensitive perspectives on alcohol use (McCaul et al., 2019): Although AUD is more prevalent in males than females (Grant et al., 2015), this gender gap is narrowing, especially in young cohorts (Keyes et al., 2011; Slade et al., 2016). This development may be related to an ongoing change of gender roles and norms via socialization across time (Slade et al., 2016). While women grow more similar to men when it comes to mere frequencies of AUD development, there are important and remarkable differences between male and female patient samples. These differences comprise biological (e.g., higher risk of cancer in women) as well as psychological variables (e.g., drinking as avoidance coping is more pronounced in men; Nolen-Hoeksema, 2004) and stress the importance of the development of gender-sensitive models and treatment approaches for AUD (see also Becker and colleagues (2016)). Moreover, a first IAT study indicates gender effects with heavy drinking females showing more negative associations towards alcohol than males (Houben et al., 2011). Expanding this research, our second aim is a deeper understanding of gender-specific alcohol associations in a clinical sample.

## MATERIALS AND METHODS

### Participants

A total of 65 patients were recruited during an alcohol-specific inpatient treatment program at 1 of 2 specialized addiction treatment centers in Switzerland (Clinic Suedhang, Bern [ $n = 49$ ], and Forel Clinic, Zurich [ $n = 13$ ]) to participate in this double-blind randomized controlled trial investigating an alcohol-specific inhibition training (Tschuempferlin et al., 2019). Only baseline and pre-intervention data were included in the present analyses. All patients fulfilled the main diagnosis of AUD according to Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; American Psychiatric Association, 2013), and were 18 to 60 years of age. At the time of the study, patients were detoxified and abstinent for a minimum of 23 days (mean: 43.29, range: 23 to 89 days). Other severe substance use disorders (except nicotine; Drug Use Identification Test; DUDIT  $\geq 25$  per substance, Berman et al., 2005) diagnosed neurocognitive disorder in the medical history, or current treatment with benzodiazepines or methylphenidate led to exclusion. Three patients had to be excluded due to insufficient IAT ( $n = 2$ ) or EEG data quality ( $n = 1$ ; see respective sections), resulting in a final sample of 62 patients.

A total of 22 controls ages 18 to 60, with nonproblematic drinking behavior (Alcohol Use Disorders Identification Test, AUDIT  $< 8$ ; Babor et al., 2001; Alcohol Use Disorder Scale, AUD-S  $< 2$ ; Wechsler et al., 1994) and low scores regarding psychopathology (Brief Symptom Check List, BSCL, GSI-t-value  $\leq 63$ ; Franke, 2000), were recruited. Current treatment for a psychiatric diagnosis and/or psychopharmacological medication, treatment for substance use disorder in the past, problematic substance use (except nicotine; DUDIT  $\geq 8$  per substance; Voluse et al., 2012), neurocognitive problems, AUD in first-degree relatives, and hearing impairments were exclusion criteria. One participant was excluded because of technical problem during the IAT, resulting in a final sample of 21 controls.

The study was approved by the ethics committees (No. 2016-00988) of the cantons of Bern and Zurich, Switzerland, in accordance with the Declaration of Helsinki and registered at ClinicalTrials.gov (NCT02968537). All subjects provided written informed consent. For more details on procedure, tasks, materials, and questionnaires used in this study, see Tschuempferlin and colleagues (2019).

### Questionnaires

The AUDIT (Babor et al., 2001), a screening instrument for AUD, and the AUD-S (adapted to DSM-5; Wechsler et al., 1994) measuring self-rated AUD symptoms, were used to assess the severity of the drinking problem. For the diagnosis of AUD, the DIA-X (adapted to DSM-5; Wittchen and Pfister, 1997) was used. The BSCL (Franke, 2000) quantified general psychopathology. Moreover, alcohol craving (Obsessive Compulsive Drinking Scale; OCDS-G; Mann and Ackermann, 2000), and positive and negative alcohol outcome (Comprehensive Alcohol Expectancy Scale, CAEQ; Demmel and Hagen, 2003; Nicolai et al., 2010) were assessed (see Table 1). A detailed overview of the sociodemographic, clinical, and motivational characteristics can be found in online Supporting Information (SI, Table S1).

### Procedure

At the time of commencing treatment, potential participants were invited to a meeting, where inclusion and exclusion criteria were checked and informed consent was obtained. During the second week of treatment program, all descriptive data, the clinical interview (DIA-X), and questionnaires (AUDIT, AUD-S) of patients were collected. At the end of the third treatment week, multichannel

EEG was assessed, while subjects completed 3 experimental tasks (see Tschuempferlin et al., 2019), including an alcohol valence IAT. Shortly before completing experimental tasks, patients filled in 2 other questionnaires (OCDS-G and CAEQ). Patients received 50 CHF (47 Euro) for the EEG measurement.

The control group first completed a screening questionnaire battery (including AUDIT, AUD-S, and BSCL) to check for inclusion and exclusion criteria. During a second appointment, controls filled in the remaining questionnaires (OCDS-G, CAEQ) before participating in EEG measurement and experimental tasks. Controls received 30 CHF (28 Euro) for their participation. Both patients and controls performed the IAT after the assessment of an alcohol-specific Go/NoGo task.

### Task and Stimuli

During the alcohol valence IAT (Greenwald et al., 1998, 2003), pictures of alcoholic or neutral beverages (target concepts) were paired with positive or negative words (affective category; see Fig. 1). For each trial, a stimulus (i.e., a picture or a word) appeared on-screen for a maximum of 1,750 ms or until an answer was given, followed by feedback lasting 200 ms. The interstimulus interval was 250 ms.

The stimuli consisted of 3 sets of 8 pictures each depicting alcoholic beverages (matched to an individual's drink of choice: beer, wine, or spirits) and 8 pictures of water. Based on previous IAT studies (e.g., Houben et al., 2011), the affective categories consisted of 8 positive (happy, jolly, energetic, funny, sociable, attractive, cheerful, and smart) and 8 negative (dull, miserable, sick, depressed, unhappy, disgusting, angry, and foolish) attributes (for German, see also Tschuempferlin et al., 2019). During "alcohol-positive" blocks, participants had to press the same response button for "alcohol" and "positive attributes," while "water" and "negative attributes" shared the other. In "alcohol-negative" blocks, "alcohol" and "negative attributes" shared 1 response button, while "water" and "positive attributes" shared the other. To minimize sequence effects depending on the starting affective category, alcohol-positive (P) and alcohol-negative blocks (N) were presented twice. In order to balance motor-related activity between both sides, the presentation order was balanced: Half the participants performed the order PNNP and the other half NPPN. Both versions were interposed with self-determined rest periods and consisted of 14 blocks.

All stimuli appeared in the middle of a white screen with the instruction to react as fast as possible by pressing 2 possible response buttons on a keyboard. Initially, and with all changes of key assignment, participants performed 2 or 3 practice blocks consisting of 16 trials each. First, they had to classify target concepts (alcohol, water) and/or the affective categories (positive, negative) to left ("a") and right ("l") by pressing a button. Then, they practiced the combination of target and affective categories. After, a longer test block of 64 trials followed. As assignment rules changed frequently, a reminder of both currently correct combinations was presented in the upper left and right corners during the whole block (see Fig. 1).

The whole task consisted of 416 trials and lasted 10 to 20 minutes, depending on the individual's speed and chosen rest periods. E-Prime v2.0 (PST, Sharpsburg, PA) was used to develop the IAT presentation and data recording.

Only the 160 alcohol trials (alcohol-positive [AP] and alcohol-negative [AN] assignments) were included in behavioral and ERP analyses, while other trials (water-positive and water-negative mappings), not representing the concept of interest, were excluded from analyses.

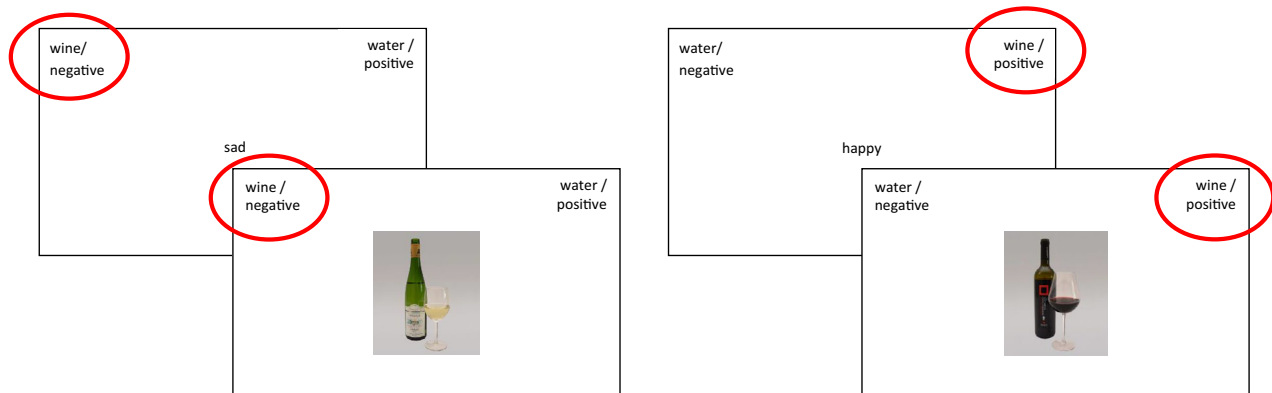
**IAT Effect:  $d_{ALC}$ .** To calculate the IAT effect (d-score  $d_{ALC}$ ), the improved scoring algorithm of Greenwald and colleagues (2003) was applied (including trials of combined practice blocks and the

**Table 1.** Descriptions and Comparison of (A) Patients and Controls and (B) Male and Female Patients

	A: Patients vs. controls					B: Male vs. female patients				
	Patients ( <i>n</i> = 62)	Controls ( <i>n</i> = 21)	<i>t</i>	<i>df</i>	<i>p</i>	Male patients ( <i>n</i> = 40)	Female patients ( <i>n</i> = 22)	<i>t</i>	<i>df</i>	<i>p</i>
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )				<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )			
Education	14.16 (2.59)	15.43 (3.14)	−1.84	81	0.070	14.03 (2.52)	14.41 (2.75)	−0.56	60	0.580
AUDIT	26.37 (5.50)	4.52 (2.16)	17.67	81	<0.001*	25.13 (4.87)	28.64 (5.97)	−2.50	60	0.015*

	Med (Range)	Med (Range)	<i>z</i>	<i>p</i>	Med (Range)	Med (Range)	<i>z</i>	<i>p</i>
Age	45 (24 to 60)	50 (25 to 58)	−0.17	0.863	44.5 (24 to 60)	46 (29 to 57)	−0.32	0.751
AUD-S	2.70 (0.00 to 4.00)	0.00 (0.00 to 0.30)	−0.68	<0.001*	2.70 (0.35 to 4.00)	2.72 (0.00 to 3.35)	−0.40	0.691
BSCl	0.90 (0.06 to 2.74)	0.15 (0.00 to 0.55)	−5.74	<0.001*	0.77 (0.06 to 2.60)	1.24 (0.15 to 2.74)	−2.71	0.007*
OCDS-G	7 (0 to 24)	2 (0 to 11)	−3.68	<0.001*	6 (0 to 21)	11 (0 to 24)	−3.00	0.003*
CAEQ 1	3.20 (0.93 to 4.60)	3.00 (1.27 to 3.93)	−1.59	0.112	3.20 (0.93 to 4.60)	3.23 (1.33 to 4.53)	−0.49	0.627
CAEQ 2	3.71 (1.00 to 5.00)	2.57 (1.14 to 4.29)	−4.56	<0.001*	3.71 (1.00 to 5.00)	3.79 (1.00 to 5.00)	−0.58	0.565
CAEQ 3	2.90 (1.00 to 4.50)	2.90 (1.30 to 4.30)	−0.49	0.622	2.95 (1.30 to 4.50)	2.85 (1.00 to 3.90)	−0.77	0.444
CAEQ 4	2.50 (0.00 to 4.75)	1.50 (1.00 to 3.75)	−3.76	<0.001*	2.50 (0.00 to 4.50)	2.63 (1.00 to 4.75)	−0.50	0.621
CAEQ 5	2.60 (0.40 to 4.60)	2.40 (1.00 to 3.80)	−1.01	0.313	2.60 (0.40 to 4.60)	2.70 (1.00 to 4.60)	−0.51	0.611

	$\chi^2$		<i>df</i>	<i>p</i>	
Gender	22f: 40m	6f: 15m	0.10	1	0.755



**Fig. 1.** Implicit Association Test (IAT). Pictures of alcohol beverages were paired with either a positive or a negative word. In alcohol-positive blocks, alcohol-related pictures and positive adjectives were assigned to the same response button, whereas in alcohol-negative blocks, alcohol-related pictures and negative adjectives shared a response button. A reminder of which picture and word type was allocated to which side was continuously presented in the upper corners of the screen. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

choice of the error penalty of 600 ms when dealing with incorrect trials), with only slight adjustments to suit the present design: All trials with RT above 300 ms from the combined blocks were used, excluding participants with more than 10% of latencies under 300 ms ( $n = 2$ ). Incorrect trials were replaced with individuals' mean RT of correct trials with 600 ms added. Then, individual means of practice and test blocks plus 1 pooled standard deviation (SD) for all practice trials as well as from all test blocks were calculated. Further, alcohol-positive trials were subtracted from alcohol-negative trials for practice and test blocks separately, normalizing these differences by the individualized pooled standard deviation of practice or test block, respectively. Computing the alcohol-related IAT effect score ( $d_{\text{ALC}}$ ), the 2 quotients were finally averaged to quantify the individual alcohol valence bias: Positive  $d_{\text{ALC}}$  scores stand for positive alcohol associations.

#### Statistical Analyses of Behavioral Data

Behavioral data were analyzed with SPSS software (IBM Corporation Armonk, NY, version 26). As  $d_{\text{ALC}}$  and RTs on correct trials

were normally distributed (Kolmogorov–Smirnov test,  $p > 0.05$  and visual inspection of histograms),  $t$ -tests were performed to analyze  $d_{\text{ALC}}$  differences between groups and RT of the correct alcohol trials was analyzed with a repeated-measures analysis of variance (ANOVA) with the factors valence (AP, AN) and group (patients, controls). The  $d_{\text{ALC}}$  was also normally distributed in both genders, and the same analysis was performed for gender differences in patients.

#### Electrophysiologic Data

**EEG Recording, Preprocessing, Data Reduction.** EEG was recorded with BrainVision Recorder (Brain Products GmbH, Gilching, Germany) using active scalp electrodes (64 positions of extended 10/10 system; sampling rate 500 Hz; band-pass filter 0.016 to 250 Hz; impedances  $\leq 20$  k $\Omega$ ; online reference FCz). Artifact rejection was performed offline with BrainVision Analyzer (Brain Products GmbH, Gilching, Germany), and eye movement artifacts were removed by means of an independent component analysis (ICA) with a plug-in incorporating the ICA algorithm



corresponding to EEGLAB ([scn.ucsd.edu/eeglab/index.php](http://scn.ucsd.edu/eeglab/index.php)); channel with excessive artifacts was interpolated using spherical splines; remaining artifacts were removed manually. Data were filtered (band-pass filter IIR [24 dB/oct]: 0.5 to 20 Hz; notch: 50Hz) and rereferenced to average reference. Individual ERPs for artifact-free and correct AP/AN assignments were computed by averaging segments from 0 to 1,000 ms poststimulus. For patients (AP: 33 to 80, AN: 26 to 80) and controls (AP: 41 to 76, AN: 32 to 78), an average of 63 artifact-free segments was included, excluding 1 person due to too few ERP trials (<20 per condition as in Stein et al., 2018). Grand-mean ERPs for AP/AN assignments and difference maps (AN minus AP) per group (patients/controls; male/female patients) were computed by averaging the individual ERPs.

**Microstate Analysis.** Using signals from all electrodes, microstate analysis segments the ERP signal data-driven and reference-independent into sequences with quasistable map topographies. These microstates stand for distinct steps in stimulus processing and represent activation of underlying networks (Koenig et al., 2013; Lehmann and Skrandies, 1980).

Microstate analyses were conducted with Ragu (Koenig et al., 2011), where a *k*-means clustering algorithm was applied to the grand-mean AP and AN ERPs of patients and controls to identify prototypical microstate maps. The optimal number of microstate maps was identified with a cross-validation procedure (Koenig et al., 2013; see SI3), and these prototypical microstate maps were then assigned to the grand-mean AP and AN ERPs of patients and controls. To examine gender effects in patients, the microstate maps observed in the patient sample were assigned to the grand-mean AP and AN ERPs of male and female patients.

To statistically test whether the factors groups (patients/controls, male/female patients), conditions (AP/AN), and their interaction had an effect on the microstate sequence, effects observed in the actual data were compared to effects obtained with data wherein the assignment of an ERP to a certain level of the factor group and/or condition was randomized (i.e., effects observed under the null hypothesis; 5,000 randomizations; (Koenig et al., 2013; see SI3)). A *p*-value of 0.05 indicates that only 5% of all effects obtained in the 5,000 randomization runs were larger than the effects obtained in our real data. Two effect parameters were extracted per microstate, factor, and condition and statistically analyzed with Ragu software: (i) *Duration* of each microstate, which is an index of time spent in a particular processing step; and (ii) *Mean Global Field Power* (GFP; Lehmann and Skrandies, 1980), which indicates overall signal strength (and thus the amount of simultaneously recruited neural resources by a particular processing step) and is independent of topography. GFP was calculated as the standard deviation of scalp voltages over all electrodes at each time point. Then, mean GFP per microstate was computed by averaging across the time points to which this microstate was assigned. As it takes 30 ms for a visual signal to reach the visual cortex (Sasaki and Watanabe, 2017), statistical analyses were restricted to microstates starting after 30 ms.

**Source Analysis.** Source analysis for significant GFP effects was computed with standardized low-resolution brain electromagnetic topography analysis (sLORETA; retrieved from <https://www.uzh.ch/keyinst/loreta.htm>; Pascual-Marqui, 2002) using log-transformed nonnormalized current density reconstruction (CDR) values, which were averaged for timeframes within a given microstate. Voxel-wise *t*-tests on averaged CDR values investigated the paired contrast (AN vs. AP/patients vs. controls/males vs. females), which had produced significant effects in the respective microstate. As statistical significance of these effects had already been determined on the scalp level, the presence of activation differences in underlying networks was considered to be statistically established. The role of sLORETA was thus not to determine statistical significance, but to identify the most probable generators

for these effects. Therefore, only a single-voxel threshold (corresponding to an alpha level of 0.5% [1-tailed]) was applied. For all clusters consisting of suprathreshold voxels, peak coordinates and cluster size are reported. In one case, a more conservative threshold (alpha level of 0.1%) had to be used due to very large effects. If analyses yielded no suprathreshold voxels, the voxel with the highest *t*-value is reported with reference to its limitation. See SI4 for more details.

## RESULTS

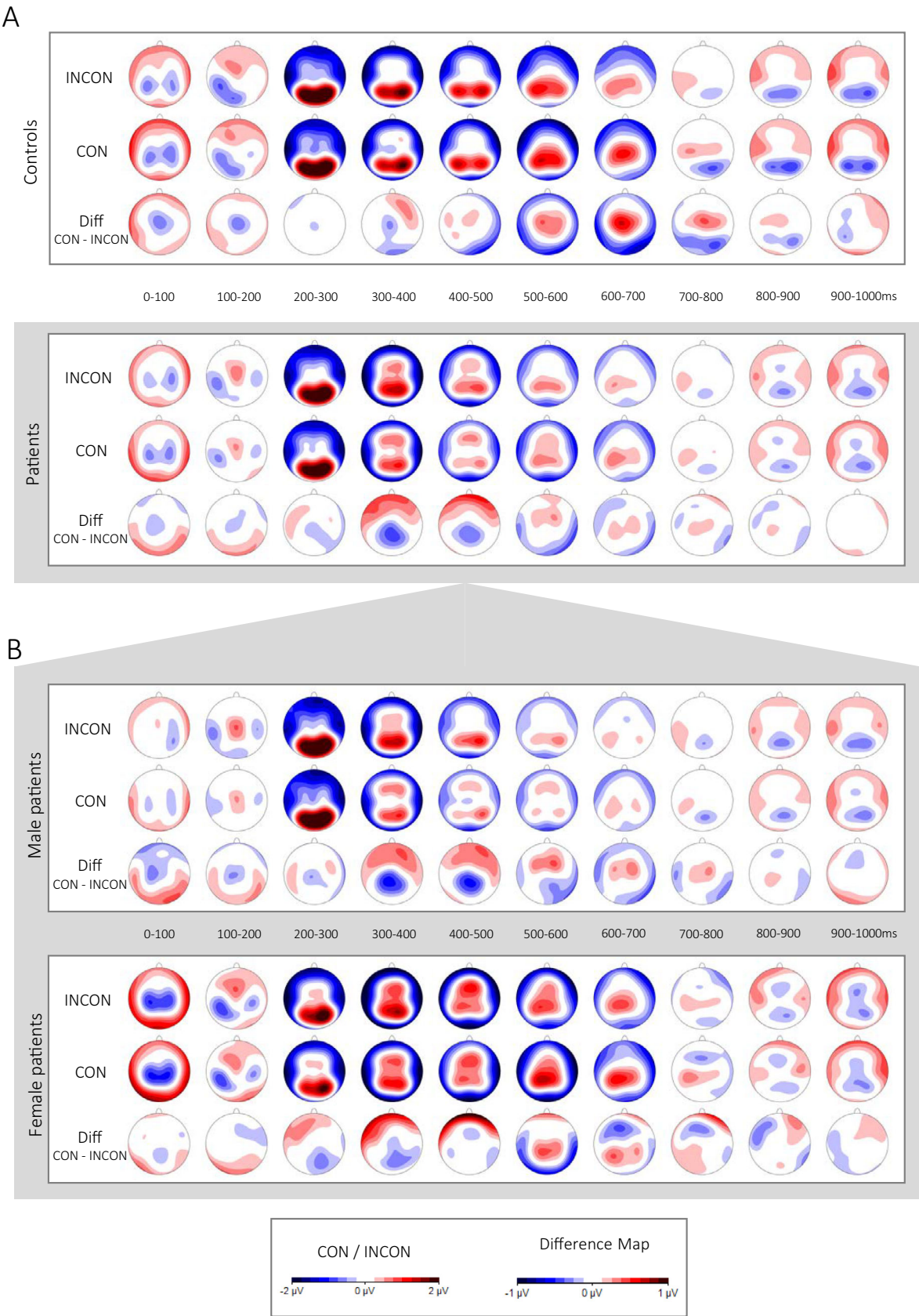
### *Implicit Alcohol Associations in AUD (Patients vs. Controls)*

**Sample Description.** Patients did not differ significantly regarding age and gender from control subjects except of a trend toward fewer years of education. As expected, patients scored higher in general psychopathology (BSCL) and alcohol-specific variables (AUDIT, AUD-S, OCDS-G, CAEQ; see Table 1). For more details, see also SI1, Table S1.

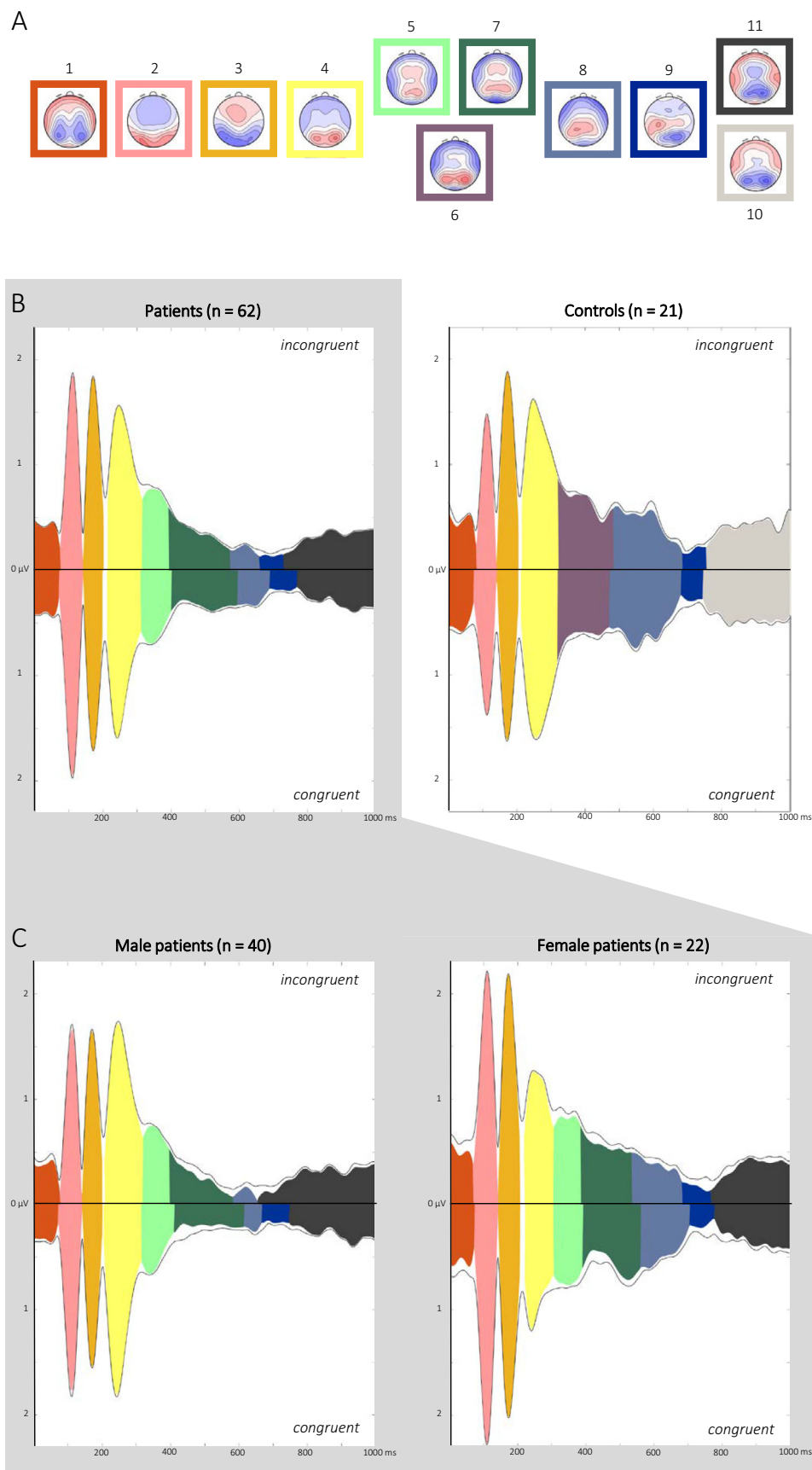
**Behavioral Data.** Mean  $d_{\text{ALC}}$  scores of patients ( $M = -0.245$ ,  $SD = 0.436$ ) and controls ( $M = -0.351$ ,  $SD = 0.416$ ) were negative, indicating negative alcohol associations and thus confirming our assumption of alcohol-negative mappings being our congruent condition. Thereby, no significant difference between groups ( $t(81) = 0.980$ ,  $p = 0.330$ ) was found. Note that the small  $d_{\text{ALC}}$  scores and large standard deviations point to a considerable heterogeneity across the sample with some scores around zero indicating potential ambivalence. Regarding error rates, no significant differences between groups were observed. Within patients, error rates were higher in the incongruent condition, while controls showed a nonsignificant trend in the same direction. While reaction times were higher in patients (compared to controls) and in the congruent (compared to incongruent) condition, there was no interaction between group and condition (see SI2.1 for details).

**ERP and Microstate Analyses.** Figure 2 shows congruent and incongruent ERPs and difference maps of patients and controls, as well as for male and female patients.

Microstate analysis revealed the optimal number of 11 microstates, which explained 93.83% of the variance in the congruent and incongruent ERPs of patients and controls. The 11 identified microstate maps and microstate sequences in the ERPs of both groups are displayed in Fig. 3A. Note that microstates 1 to 4 and 8 and 9 occur in both groups, while microstates 5, 7, and 11 occur only in patients and microstates 6 and 10 occur only in controls. Analyses for these group-specific microstates were conducted separately for patients or controls, whereby no main effect of group can be reported. Analyses of all other microstates were conducted in the total sample ( $N = 83$ ). Note that only significant results or trends are reported in the text. If not mentioned, there was no significant effect. Microstate occurrence per group and valence are depicted in SI3 (Table S2).



**Fig. 2.** ERP topographies from 0 to 1,000 ms after stimulus presentation for **(A)** patients and controls and **(B)** male and female patients. *CON*, congruent (alcohol-negative mappings); *INCON*, incongruent (alcohol-positive mappings); *Diff* CON – INCON, difference maps congruent minus incongruent. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]





**Fig. 3.** Microstate analyses. **(A)** 11 microstate maps computed for the congruent and incongruent ERPs of patients and controls. Microstates occurring in both groups are depicted in the central row, patient-specific (**upper row**), and control-specific microstates (**lower row**). **(B)** Microstate assignment to the ERPs of patients and controls. The assignment of a microstate to a specific time point is indicated by color-coding depicted under the respective GFP curve. The precise timeframes of occurrence per group and condition are indicated in Tables S2 and S3. The y-axis indicates that the GFP curve of the incongruent ERP (alcohol-positive mappings) is plotted with positive values up, while the congruent ERP (alcohol-negative mappings) is flipped and plotted with positive values down. **(C)** Microstate assignment to the ERPs of male and female patients indicated by color codes under the respective GFP curve. Note again that the GFP curve of the incongruent ERP is plotted with positive values up and the congruent ERP with positive values down. The x-axis represents time (ms) after stimulus presentation; the y-axis refers to the global field power, which is displayed in microvolts ( $\mu\text{V}$ ). [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

*Early Microstates (2 to 4: Patients and Controls)-approximately 0 to 320 ms.* In microstate 3, significantly higher map strength occurred in incongruent than in congruent trials ( $1.30 \mu\text{V}$  vs.  $1.20 \mu\text{V}$ ,  $p = 0.003$ ).

*Intermediate Microstates (5/7: Patients; 6: Controls; 8: Patients and Controls)-approximately 320 to 700 ms.* Between 300 and 400 ms, patients displayed microstate 5, with a significant effect of valence on map strength: Incongruent trials showed higher GFP than congruent ones ( $0.77 \mu\text{V}$  vs.  $0.64 \mu\text{V}$ ,  $p = 0.002$ ). During this timeframe, controls exhibited microstate 6, where no significant effect was found ( $\text{GFP} = 0.67 \mu\text{V}$  in both conditions,  $p = 0.94$ ). In microstate 8, controls had significantly higher GFP than patients ( $0.55 \mu\text{V}$  vs.  $0.28 \mu\text{V}$ ,  $p = 0.007$ ).

*Late Microstates (9: Patients and Controls; 10: Controls; 11: Patients)-approximately 700 to 1,000 ms.* A significant main effect of valence as well as a group-by-valence interaction was observed in microstate 9: More GFP was elicited during congruent than incongruent trials in the whole sample ( $0.23 \mu\text{V}$  vs.  $0.18 \mu\text{V}$ ,  $p = 0.004$ ), but this difference was larger in controls (patients:  $0.21 \mu\text{V}$  vs.  $0.18 \mu\text{V}$ , controls:  $0.34 \mu\text{V}$  vs.  $0.22 \mu\text{V}$ , interaction effect:  $p = 0.049$ ). Microstate 10 was unique for controls and microstate 11 for patients, respectively.

*Source Analyses.* To estimate brain regions generating significant main effects and interactions regarding GFP, sLORETA source analysis was applied (Table 2a; Fig. 4A).

For the main effect of valence in microstate 3, with higher GFP for incongruent than congruent trials, the left superior frontal gyrus was identified as generator. A valence effect in patient-specific microstate 5 revealed higher activation in incongruent compared to congruent trials in clusters comprising the right inferior frontal gyrus, left superior temporal gyrus, right insula, and inferior parietal gyrus, as well as in bilateral posterior cingulate cortex (PCC, left also spreading into parahippocampal gyrus). sLORETA of the group effect in microstate 8 yielded no suprathreshold voxels; thus, the estimation of higher precuneus activation in controls than patients must be interpreted with caution. Source estimation of the main effect of valence, as well as the group-by-valence interaction in microstate 9, indicated that higher superior frontal gyrus activation during AN trials was driven by patients. The controls rather displayed higher activation in

bilateral inferior parietal gyrus as well as the left precuneus, insula, and middle frontal gyrus.

#### Gender Effects in Patients

*Sample Description.* Male and female patients did not differ significantly in age and education. However, women displayed higher scores in the screening instrument for AUD (AUDIT), but not in the number of criteria met (AUD-S). Moreover, self-reported scores in general psychopathology (BSCL) and craving were significantly higher in females than in males (see Table 1).

*Behavioral Data.* Both, male and female patients, showed negative alcohol associations ( $d_{\text{ALC}}$  scores; men:  $M = -0.294$ ,  $SD = 0.065$ ; women:  $M = -0.1556$ ,  $SD = 0.101$ ), which did not differ significantly ( $t(60) = -1.179$ ,  $p = 0.236$ ). Men displayed higher error rates than women, and both genders made more errors in incongruent than congruent trials (see SI2).

*ERP and Microstate Analyses.* For the 9 microstates occurring in patients, analyses of gender effects were conducted. The microstate sequences for ERPs of male ( $n = 41$ ) and female ( $n = 21$ ) patients are displayed in Fig. 3C.

The 9 microstates explained 87.30% of the variance in the congruent and incongruent ERPs of both groups. As main effects of valence were already reported in the previous section, only significant (or trends in) main effects of group and interactions (valence  $\times$  group) in duration and mean GFP are described here. For detailed information of the time windows of each microstate in both groups and conditions, please see Table S3 in SI3. Note that the respective analyses were also performed in controls (see SI5). These yielded 1 single effect and will only be referred to determine whether observed effects selectively occurred in patients or must rather be interpreted as general gender effects.

*Early Microstates (2 to 4)-approximately 0 to 320 ms.* No significant gender effect was found in early microstates.

*Intermediate Microstates (5, 7, and 8)-approximately 320 to 700 ms.* During microstate 5, no significant effect was observed. A main effect of gender was found in microstates 7 and 8: Women displayed significantly more GFP than men



(microstate 7: 0.62  $\mu\text{V}$  vs. 0.27  $\mu\text{V}$ ,  $p < 0.001$ ; microstate 8: 0.46  $\mu\text{V}$  vs. 0.21  $\mu\text{V}$ ,  $p = 0.003$ ). Note that no significant effect in this timeframe, and rather a reversed descriptive pattern, was observed in the respective analyses in controls (see SI5; both  $p$ -values  $> 0.4$ ).

*Late Microstates (9 and 11) - approximately 700 to 1,000 ms.* The pattern of appearance of microstate 9 resembled an interaction effect, as it was visible in both ERPs of females, but only in the congruent ERP of males. However, as it did not occur in all conditions, the interaction could not be statistically tested for the parameter of mean GFP and did not reach significance for the parameter of duration. Note that the respective analysis in controls also pointed toward an interaction effect; thus, this seemed to be not restricted to a patient sample.

*Source Analyses.* Localization of significant gender effects with sLORETA indicated that female patients displayed higher activation than male patients in 2 big clusters in left cuneus/precuneus/PPC and right precuneus/cuneus/PCC during microstate 7, while in microstate 8, no suprathreshold voxels were observed (Table 2b; Fig. 4B).

## DISCUSSION

This study investigates neurophysiological correlates of implicit associations toward alcohol in patients with AUD (first aim) for the first time and explores potential gender effects regarding these processes (second aim).

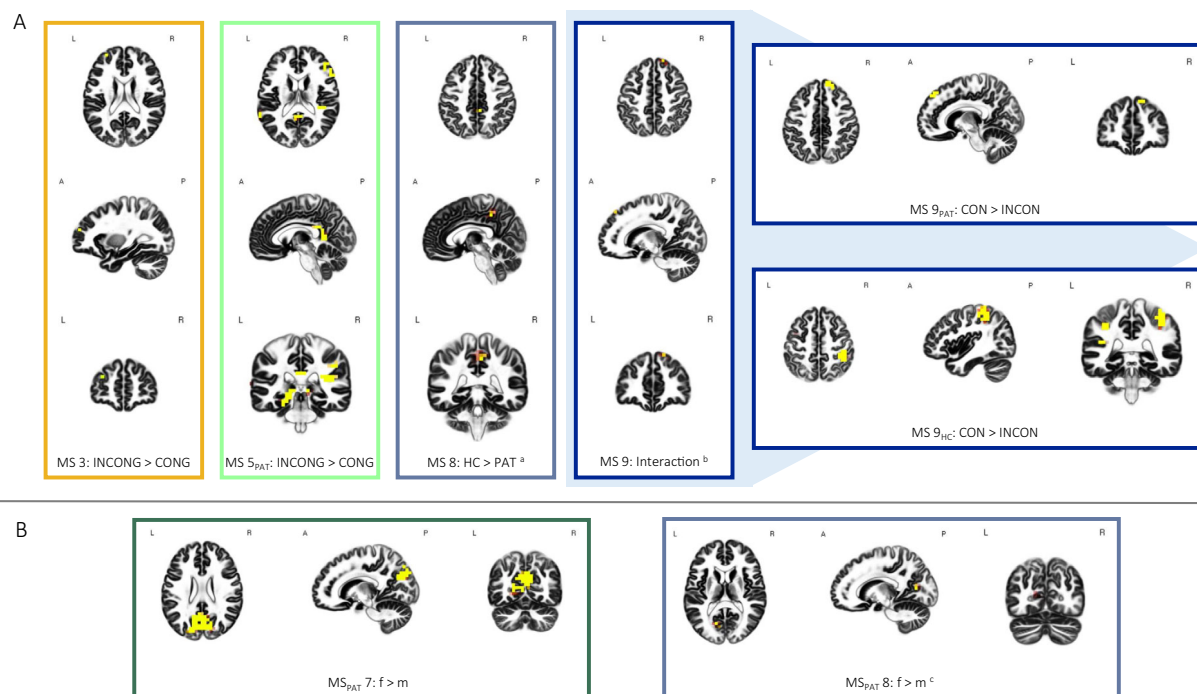
### Neurophysiology of Implicit Alcohol Associations in AUD

Both groups showed negative  $d_{\text{ALC}}$  scores, validating that alcohol-negative trials reflect the congruent condition. While no behavioral group differences occurred, groups differed significantly on a neurophysiological level. From 320 ms onwards, ERPs of patients and controls were characterized by distinct topographic patterns, with 2 specific topographies (microstates 5 and 7) in patients, and another distinct topography (microstate 6) in controls. Differing from controls' topography, and also from findings in other healthy samples during an IAT (O'Toole and Barnes-Holmes, 2009; Schiller et al., 2016; Williams and Themanson, 2011), patients displayed an extended frontal positivity during this timeframe (Figs 2 and 3). Interestingly, ERPs varied between conditions during this timeframe only in patients with microstate 5

**Table 2.** sLORETA Localization of the Timeframe With Significant Effects per Microstate in (A) Patients and Controls and (B) Male and Female Patients

Microstates Structures per effect	R/L	BA	MNI coordinates			Cluster size ( <i>n</i> voxel)	<i>t</i> -value
			<i>x</i>	<i>y</i>	<i>z</i>		
A Patients and controls							
MS 3 (142 to 200 ms): Valence AP> AN							
Superior frontal gyrus	L	10	−30	55	20	1	−2.64
MS 5 (316 to 394 ms): Valence AP > AN in Pat							
Inferior frontal gyrus	R	46	55	30	15	14	−3.19
Superior temporal gyrus	L	22	−65	−50	10	7	−3.19
Posterior cingulate (parahippocampal gyrus)	L	30 (27)	−10	−40	0	33 (12)	−3.14
Insula	R	13	40	−35	20	8	−3.00
Inferior parietal gyrus	R	40	55	−40	35	9	−2.97
Posterior cingulate	R	30	0	−45	20	18	−2.92
MS 8 (586 to 674 ms): Group HC > Pat							
Precuneus <sup>a</sup>	R	7	5	−40	45	1	−1.50
MS 9 (690 to 728 ms): Valence AN > AP							
Superior frontal gyrus	R	8	5	45	50	3	2.76
MS 9 (690 to 728 ms): Group × valence							
Superior frontal gyrus <sup>b</sup>	R	8	15	45	50	1	1.83
Pat: Valence AN > AP							
Superior frontal gyrus	R	8	20	40	50	12	3.22
HC: Valence AN > AP							
Inferior parietal gyrus	L	40	−35	−40	40	12	4.84
Inferior parietal gyrus	R	40	45	−40	55	29	3.58
Precuneus	L	31	−15	−35	40	7	3.21
Insula	L	13	−45	−40	20	2	3.05
Middle frontal gyrus	L	6	−35	5	50	1	2.87
B Male and female patients							
MS 7 (398 to 517 ms): Group female > male							
Cuneus/precuneus/posterior cingulate	L	31/31/18	−10	−70	15	109 (47/37/25)	−4.63
Precuneus/cuneus/posterior cingulate	R	31/19/23	5	−70	20	78 (47/26/5)	−4.24
MS 8 (616 to 654 ms): Group female > male							
Cuneus <sup>c</sup>	L	30	−15	−70	10	1	−2.20

*N*, alcohol-negative (congruent) mappings; *AP*, alcohol-positive (incongruent) mappings; *BA*, Brodmann's area; *HC*, healthy control; *MS*, microstate; *MNI*, Montreal Neurological Institute; *Pat*, patients. Timeframes of microstate reflect the maximal onset and minimal offset of all conditions. Note: Clusters are labeled according to the regions with the dominant peak voxels. Superscripted letters indicate that there is only a statistical trend corresponding to <sup>a</sup>  $t_{81} = 1.50$ , <sup>b</sup>  $t_{81} = 1.83$ , and <sup>c</sup>  $t_{60} = -2.20$ . MS 5 was patient-specific and did not occur in controls.



**Fig. 4.** sLORETA localizations. **(A)** Source estimations of significant effects in the analyses of patients and controls. Note that microstate 5 is patient-specific, and this valence effect occurred thus only in patients. **(B)** Source estimations of significant effects in the analyses of male and female patients. A, anterior; P, posterior; L, left; R, right; MS, microstate; CON, congruent (alcohol-negative mappings); INCON, incongruent (alcohol-positive mappings); Pat, patients; HC, healthy controls; m, male; f, female. Note that in general, scaling was adjusted to correspond to an alpha level of 0.5% (1-tailed). In microstate 7 (part B), a more conservative threshold (alpha level of 0.1%) had to be used due to very large effects. Also, superscripted letters indicate that there was only a statistical trend corresponding to <sup>a</sup>  $t_{61} = 1.50$ , <sup>b</sup>  $t_{61} = 1.83$ , and <sup>c</sup>  $t_{60} = -2.20$ . For the specifications of the involved brain regions and the specific timeframes used in the analyses, see Table 2. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

displaying higher signal strength during incongruent trials, whereas controls did not differ between conditions. Given the timing and generators of this effect, this finding could represent additional resource allocation to semantic and self-referential processing, as well as to adaptive task management in patients. Generators of this effect in the patient sample included the bilateral PCC (BA 33; including left parahippocampal gyrus, BA 12), the right inferior frontal gyrus (rIFG, BA 46), the right inferior parietal gyrus (rIPG, BA 40), and the left superior temporal gyrus (ISTG, BA 22). The PCC has a key function in the default mode network (Raichle et al., 2001) and has generally been related to enhanced recruitment of additional resources during difficult and emotional tasks (Maddock, 1999; Pearson et al., 2009), especially when adaptations to the current model of the world are necessary (i.e., congruent to incongruent condition; Pearson et al., 2011). The left PCC has further been proposed to be an interface between episodic encoding and semantic retrieval (Binder et al., 2009), while the ISTG (BA 22) is known to react to semantic violations (Friederici et al., 2003). Given that ERP effects in this timeframe have been repeatedly associated with semantic processing (e.g., Kutas and Hillyard, 1980), semantic integration might evolve differently and more effortful on incongruent trials in our patient sample. Enhanced activity in the IFG has previously been found in incongruent trials of an IAT (Knutson et al., 2006)

and is assumed to be linked to inhibition (Aron et al., 2014; Everitt and Robbins, 2016). Extending this interpretation, and as patients also made more errors in incongruent trials, their increased rIPG activity could reflect increased allocation of attentional resources to these subjectively more difficult trials (Rubia et al., 2003). Another source was the right insula (BA 13), which has been implied in addition in general (Volkow and Baler, 2014) and also specifically in alcohol-related implicit associations of heavy drinkers (Ames et al., 2014). Insula activation is associated with self-referential processing, interoception, and self-awareness (Craig, 2009). Taken together, in patients, the higher activation (indicated by higher GFP in microstate 5) of different brain regions during the incongruent compared to the congruent condition might represent additional resource allocation to semantic and self-referential processing as well as adaptive task management. Possibly, these additional resources are necessary during the processing of positive alcohol associations (incongruent condition), because in these recently abstinent inpatients, alcohol is an issue which is related to highly personal and ambivalent topics.

At approximately 600 ms, when both groups converged again onto the same microstates, our results might indicate enhanced emotional processing in controls, especially during the congruent condition. During microstate 8, controls showed higher signal strength. In microstate 9, a significant

group  $\times$  valence interaction revealed higher signal strength in congruent trials, which is more pronounced in controls. Both, microstate 8 (586 to 674 ms) and 9 (690 to 728 ms) occurred in the typical LPP timeframe, which is thought to index processing of emotional stimuli (Cuthbert et al., 2000; Dillon et al., 2006). Difference maps (congruent minus incongruent) of both timeframes indicate typical LPP maps in controls, but not in patients (see Fig. 2). Similarly, earlier IAT studies in healthy subjects also reported larger LPP amplitudes in congruent trials (Hurtado et al., 2009; O'Toole and Barnes-Holmes, 2009; Williams and Themanson, 2011), which was interpreted as stronger emotional activation in congruent than incongruent trials. The LPP enhancement in microstate 8 in controls (in contrast to patients) might therefore be indicative of enhanced emotional processing during this timeframe.

Localization of the interaction between group and condition in microstate 9 indicated that in patients, activity in the right superior frontal gyrus (rSFG; BA 8) differed between conditions, while controls showed more activation in inferior parietal gyrus bilaterally (IPG; BA 40), the left precuneus (BA 7), and the insula (BA 2) during congruent as compared to incongruent trials. One could conclude that in the late LPP timeframe, controls might invest more in emotional processing during congruent compared to incongruent trials, whereas patients exhibit primarily prefrontal processes, possibly reflecting motor preparation (average reaction time of congruent (711 ms;  $\pm 81$  ms) and incongruent (756 ms;  $\pm 97$  ms) trials lies in the timeframe of microstate 9). The patients' activation difference in SFG is in line with previous findings (Egenolf et al., 2013), showing higher SFG activation during the LPP timeframe in congruent conditions, possibly reflecting an earlier start in response preparation in congruent as compared to incongruent trials, as it occurred shortly before the mean RT in congruent trials. Considering activation differences in controls, the IPG, previously linked to conscious motor intention (Desmurget and Sirigu, 2012), might reflect motor response preparation as well, but activation of the precuneus (Ferri et al., 2016; Maddock, 1999) and insula (Gasquoine, 2014) are more indicative of enhanced emotional and self-reflective processing during congruent as compared to incongruent trials.

The only effect, which was similar in patients and controls, occurred during an early microstate (microstate 3, 142 to 200 ms) with a main effect of valence indicating that both groups had higher map strength in incongruent trials, probably reflecting greater levels of response conflict during the processing of incongruent mappings (Folstein and Van Petten, 2008; Yeung et al., 2004). The effect was localized in the left superior frontal gyrus (BA 10), showing higher activation during incongruent than congruent trials. This finding is in line with earlier IAT studies interpreting larger ERP amplitudes (Coates and Campbell, 2010; Xiao et al., 2015) and enhanced prefrontal activity (Ames et al., 2014; Chee et al., 2000) during incongruent trials as indicators of enhanced response conflict.

Regarding behavioral data, this is the first study comparing bipolar IAT results between patients and controls. Building on reports of heavy drinkers having less negative  $d_{\text{ALC}}$  scores than light drinkers (Wiers et al., 2005; Wiers et al., 2002), we expected negative associations toward alcohol in patients as well as controls, but hypothesized that this bias would be more pronounced in controls. With a negative bias in both groups and more negative scores in controls, behavioral data ( $d_{\text{ALC}}$ ) descriptively aligned with this pattern, but there was no significant group difference. While the negative  $d_{\text{ALC}}$  scores are in line with an earlier IAT study in AUD (De Houwer et al., 2004), the absence of group effects was unexpected. Perhaps, the fact that we investigated patients currently undergoing intensive AUD treatment (see also limitations), which might have shifted their implicit bias to become more negative, obscured potential group differences. To test this interpretation, future research could investigate IAT effects in individuals with AUD, who are not undergoing treatment.

### *Gender Effects in Patients*

The results of the gender-specific behavioral analysis revealed no  $d_{\text{ALC}}$  score differences. On the neurophysiological level, gender effects emerged. The ERP signal of female patients, with higher LPP signal strength and overlapping generator networks, seems to resemble the controls' ERP more strongly than that of male patients: Between 400 and 650 ms after stimulus presentation (microstates 7 and 8) females showed higher signal strength than males. Generators for this effect were estimated in 2 big bilateral clusters including the cuneus (BA 31), the precuneus (BA 19, 31), and the PCC (BA 23, 18). Thus, not only the topographies of female patients during both microstates (398 to 654 ms) resembled the topographies of the controls (Fig. 2) and might be interpreted in terms of slightly frontalized LPP maps. Moreover, the localization of this gender effect with higher activation in women than in men partially matches the finding of higher precuneus activation in controls than in patients. Both the PCC (Maddock et al., 2003) and the precuneus (Ferri et al., 2016) have previously been associated with emotional processing, which thus might be attenuated selectively in male patients. Note that similar analyses in controls yielded no such effects; therefore, general gender effects on signal strength during this timeframe, as previously reported (i.e., Syrjanen and Wiens, 2013), are unlikely. Taking theoretical considerations into account (Becker et al., 2016), an interplay of psychosocial (e.g., more social sanctions for women due to alcohol consumption, e.g., de Visser and McDonnell, 2012, or differential drinking patterns, e.g., Erol and Karpyak, 2015) and biological factors (e.g., sex differences in genetics, hormones, or alcohol sensitivity) might underlie the observed gender effects in the neurophysiology of implicit alcohol cognitions in patients. Regarding psychosocial aspects, it is possible that differences in sociocultural experiences (as, e.g., the experience of social sanctions;



de Visser and McDonnell, 2012) foster the development of differential processing of implicit alcohol associations. Regarding biological aspects, a heightened alcohol sensitivity might lead to female patients implicitly assessing alcohol differently from the less sensitive men (Nolen-Hoeksema, 2004). The fact that the gender effects were not observed in the healthy control sample suggests that more general differences (like hormones or genes) are less likely to underlie these effects, unless they interact with alcohol consumption.

In summary, extending earlier reports regarding gender differences with respect to the development and maintenance of AUD (Nolen-Hoeksema, 2004), we report differences in the neurophysiological basis of implicit alcohol associations. Consequently, the consideration of gender effects in future research and the transferral of such findings into the development of gender-sensitive treatment seem of great importance.

### Limitations

Patients included in this study attended a specialized inpatient treatment program for AUD, which might have influenced their strategies and state of mind during IAT participation. During AUD treatment, patients likely focused on negative aspects of alcohol use. In the context of the IAT performance, some expressed concerns about associating the combination of alcohol and positive attributes. Although our data indicate negative implicit alcohol associations in patients, which is in line with previous clinical studies (De Houwer et al., 2004; Dickson et al., 2013), high variability in  $d_{\text{ALC}}$  scores reveals a considerable heterogeneity, which may be associated with ambivalence. This ambivalence is also reflected in research with unipolar IATs, which test either positive or negative associations separately against neutral ones. Here, both positive and negative associations are measurable in patients (Dickson et al., 2013). Future research in in- and outpatient settings should investigate the role of possible applied strategies and address the issue of considerable ambivalence in patient samples.

One might argue that the reported effects could be explained by differences in variables like age, psychopathology, drinking history, craving, or alcohol expectancies. However, this possibility was tested in additional analyses (see also SI3.3), none of which yielded significant results. It is therefore unlikely that our effects are better explained by these variables. A further limitation lies in the a priori assumptions, the limited spatial resolution, and failure/scarcely accountability to consider for individual brain morphology introduced by sLORETA (Luck, 2005).

### CONCLUSIONS

Our results improve the understanding of implicit alcohol associations in patients with AUD, thus broadening our knowledge about implicit (unconscious) processes relevant to AUD. Although not discernible on a behavioral level, patients display distinct mental processing, as indicated on

the neurophysiological level by divergent microstates and alterations in microstate signal strength. We interpret these neurophysiological differences between recently abstinent patients with AUD currently attending a specialized inpatient program to suggest a: (i) greater effort regarding semantic and self-relevant integration around 400 ms during incongruent trials and (ii) attenuated emotional processing during the LPP timeframe compared to controls. Interestingly, this emotional blunting was reduced in female patients, who showed more similarities with controls than male patients. Extending research on explicit cognitions (e.g., Nolen-Hoeksema, 2004), our results thus show that implicit alcohol cognitions also vary with gender. These effects might be a result of a complex interplay between biology, the physical environment, and psychosocial factors resulting in diverse biological processes of addiction (Becker et al., 2016). As the gender gap in drinking patterns and AUD development narrows and more treatment-seeking female patients are to be expected (Grant et al., 2015; Keyes et al., 2011), such neurophysiological results highlight the importance of gender-sensitive research and encourage the development of gender-sensitive treatment approaches to AUD.

### ACKNOWLEDGMENTS

We thank Nadja Zimmermann, Amanda Schnyder, Sarah Diezig, Plinio Tettamanti, Nadine Rennhard, Benjamin Erb and Sara Lustenberger for their valuable help in data collection and/or preprocessing. We also thank all other involved students for their engagement in the recruitment and data collection. Greatest thanks goes to all participants!

### FUNDING

This project was supported by a grant provided by the Swiss National Science Foundation (SNSF) (No: 105319\_159286) to MS, LS, and FM.

### CONFLICT OF INTEREST

The authors declare no conflict of interest.

### AUTHOR CONTRIBUTIONS

MS, LS, and FM designed the study. HB and RT collected data. SR, AK, and PP provided the infrastructure and access during recruitment. RT, MS, and TK analyzed and interpreted the data with the support of LS, HB, and FM. RT wrote the first draft, and all authors reviewed and approved the final manuscript.

### REFERENCES

- American Psychiatric Association (2013) Diagnostic and Statistical Manual of Mental Disorders: Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association, Arlington, VA.

- Ames SL, Grenard JL, He Q, Stacy AW, Wong SW, Xiao L, Xue G, Bechara A (2014) Functional imaging of an alcohol-Implicit Association Test (IAT). *Addict Biol* 19:467–481.
- Ames SL, Grenard JL, Stacy AW, Xiao L, He Q, Wong SW, Xue G, Wiers RW, Bechara A (2013) Functional imaging of implicit marijuana associations during performance on an Implicit Association Test (IAT). *Behav Brain Res* 256:494–502.
- Aron AR, Robbins TW, Poldrack RA (2014) Inhibition and the right inferior frontal cortex: one decade on. *Trends Cogn Sci* 18:177–185.
- Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG (2001) AUDIT. The Alcohol Use Disorders Identification Test. Guidelines for Use in Primary Care. World Health Organization, Geneva.
- Becker JB, McClellan M, Reed BG (2016) Sociocultural context for sex differences in addiction. *Addict Biol* 21:1052–1059.
- Berman AH, Bergman H, Palmstierna T, Schlyter F (2005) Evaluation of the Drug Use Disorders Identification Test (DUDIT) in criminal justice and detoxification settings and in a Swedish population sample. *Eur Addict Res* 11:22–31.
- Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex* 19:2767–2796.
- Chee MWL, Sriram N, Soon CS, Lee KM (2000) Dorsolateral prefrontal cortex and the implicit association of concepts and attributes. *NeuroReport* 11:135–140.
- Chen L, Zhou H, Gu Y, Wang S, Wang J, Tian L, Zhu H, Zhou Z (2018) The neural correlates of implicit cognitive bias toward internet-related cues in internet addiction: an ERP study. *Front Psychiatry* 9:421.
- Coates MA, Campbell KB (2010) Event-related potential measures of processing during an Implicit Association Test. *NeuroReport* 21:1029–1033.
- Craig AD (2009) How do you feel now? The anterior insula and human awareness. *Nat Rev Neurosci* 10:59–70.
- Cuthbert BN, Schupp HT, Bradley MM, Birbaumer N, Lang PJ (2000) Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biol Psychol* 52:95–111.
- De Houwer J, Crombez G, Koster EH, De Beul N (2004) Implicit alcohol-related cognitions in a clinical sample of heavy drinkers. *J Behav Ther Exp Psychiatry* 35:275–286.
- de Visser RO, McDonnell EJ (2012) 'That's OK. He's a guy': a mixed-methods study of gender double-standards for alcohol use. *Psychol Health* 27:618–639.
- Demmel R, Hagen J (2003) The comprehensive alcohol expectancy questionnaire: I. Scale development. *SUCHT* 49:292–299.
- Desmurget M, Sirigu A (2012) Conscious motor intention emerges in the inferior parietal lobule. *Curr Opin Neurobiol* 22:1004–1011.
- Dickson JM, Gately C, Field M (2013) Alcohol dependent patients have weak negative rather than strong positive implicit alcohol associations. *Psychopharmacology* 228:603–610.
- Dillon DG, Cooper JJ, Grent-'t-Jong T, Woldorff MG, LaBar KS (2006) Dissociation of event-related potentials indexing arousal and semantic cohesion during emotional word encoding. *Brain Cogn* 62:43–57.
- Egenolf Y, Stein M, Koenig T, Grosse Holtforth M, Dierks T, Caspar F (2013) Tracking the implicit self using event-related potentials. *Cogn Affect Behav Neurosci* 13:885–899.
- Erol A, Karpyak VM (2015) Sex and gender-related differences in alcohol use and its consequences: contemporary knowledge and future research considerations. *Drug Alcohol Depend* 156:1–13.
- Everitt BJ, Robbins TW (2016) Drug addiction: updating actions to habits to compulsions ten years on. *Annu Rev Psychol* 67:23–50.
- Ferri J, Schmidt J, Hajcak G, Canli T (2016) Emotion regulation and amygdala-precuneus connectivity: focusing on attentional deployment. *Cogn Affect Behav Neurosci* 16:991–1002.
- Fleischhauer M, Strobel A, Diers K, Enge S (2014) Electrophysiological evidence for early perceptual facilitation and efficient categorization of self-related stimuli during an Implicit Association Test measuring neuroticism. *Psychophysiology* 51:142–151.
- Folstein JR, Van Petten C (2008) Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* 45:152–170.
- Franke GH (2000) Brief Symptom Inventory von L.R. Derogatis—Deutsches Manual. [Brief symptom Inventory of L.R. Derogatis—German Manual]. Beltz, Göttingen.
- Friederici AD, Rueschemeyer SA, Hahne A, Fiebach CJ (2003) The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cereb Cortex* 13:170–177.
- Gasquoine PG (2014) Contributions of the insula to cognition and emotion. *Neuropsychol Rev* 24:77–87.
- Grant BF, Goldstein RB, Saha TD, Chou SP, Jung J, Zhang H, Pickering RP, Ruan WJ, Smith SM, Huang B, Hasin DS (2015) Epidemiology of DSM-5 alcohol use disorder: results from the National Epidemiologic Survey on Alcohol and Related Conditions III. *JAMA Psychiatry* 72:757–766.
- Greenwald AG, McGhee DE, Schwartz JLK (1998) Measuring individual differences in implicit cognition: the Implicit Association Test. *JPS* 74:1464–1480.
- Greenwald AG, Nosek BA, Banaji MR (2003) Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *J Person Soc Psychol* 85:197–216.
- Houben K, Nederkoorn C, Wiers RW, Jansen A (2011) Resisting temptation: decreasing alcohol-related affect and drinking behavior by training response inhibition. *Drug Alcohol Depend* 116:132–136.
- Houben K, Wiers RW (2008) Implicitly positive about alcohol? Implicit positive associations predict drinking behavior. *Addict Behav* 33:979–986.
- Hurtado E, Haye A, Gonzalez R, Manes F, Ibanez A (2009) Contextual blending of ingroup/outgroup face stimuli and word valence: LPP modulation and convergence of measures. *BMC Neurosci* 10:69.
- Jajodia A, Earleywine M (2003) Measuring alcohol expectancies with the implicit association test. *Psychol Addict Behav* 17:126–133.
- Keyes KM, Li G, Hasin DS (2011) Birth cohort effects and gender differences in alcohol epidemiology: a review and synthesis. *Alcohol Clin Exp Res* 35:2101–2112.
- Knutson KM, Mah L, Manly CF, Grafman J (2007) Neural correlates of automatic beliefs about gender and race. *Hum Brain Mapp* 28:915–930.
- Knutson KM, Wood JN, Spampinato MV, Grafman J (2006) Politics on the brain: an fMRI investigation. *Soc Neurosci* 1:25–40.
- Koenig T, Kottlow M, Stein M, Melie-Garcia L (2011) Ragú: a free tool for the analysis of EEG and MEG event-related scalp field data using global randomization statistics. *Comput Intell Neurosci* 4:1–14.
- Koenig T, Stein M, Grieder M, Kottlow M (2013) A tutorial on data-driven methods for statistically assessing ERP topographies. *Brain Topogr* 27:72–83.
- Kutas M, Hillyard SA (1980) Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207:203–205.
- Lehmann D, Skrandies W (1980) Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr Clin Neurophysiol* 48:609–621.
- Luck SJ (2005) ERP localization, in *An Introduction to the Event-Related Potential Technique*, 1st ed. (Luck SJ ed), pp 267–300. MIT Press, Cambridge, MA.
- Luo Q, Nakic M, Wheatley T, Richell R, Martin A, Blair RJ (2006) The neural basis of implicit moral attitude—an IAT study using event-related fMRI. *NeuroImage* 30:1449–1457.
- Maddock RJ (1999) The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. *Trends Neurosci* 22:310–316.
- Maddock RJ, Garrett AS, Buonocore MH (2003) Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. *Hum Brain Mapp* 18:30–41.
- Mann K, Ackermann K (2000) Die OCDS-G: Psychometrische Kennwerte der deutschen Version der Obsessive Compulsive Drinking Scale. *SUCHT* 46:90–100.

- McCaul ME, Roach D, Hasin DS, Weisner C, Chang G, Sinha R (2019) Alcohol and women: a brief overview. *Alcohol Clin Exp Res* 43:774–779.
- Nicolai J, Demmel R, Moshagen M (2010) The comprehensive alcohol expectancy questionnaire: confirmatory factor analysis, scale refinement, and further validation. *J Pers Assess* 92:400–409.
- Nolen-Hoeksema S (2004) Gender differences in risk factors and consequences for alcohol use and problems. *Clin Psychol Rev* 24:981–1010.
- O'Toole C, Barnes-Holmes D (2009) Electrophysiological activity generated during the implicit association test: a study using event-related potentials. *Psychol Rec* 59:207–219.
- Pascual-Marqui RD (2002) Standardized low resolution brain electromagnetic tomography (sLORETA): technical details. *Meth Find Exp Clin Pharmacol* 24:5–12.
- Pearson JM, Hayden BY, Raghavachari S, Platt ML (2009) Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. *Curr Biol* 19:1532–1537.
- Pearson JM, Heilbronner SR, Barack DL, Hayden BY, Platt ML (2011) Posterior cingulate cortex: adapting behavior to a changing world. *Trends Cogn Sci* 15:143–151.
- Raichle ME, Macleod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. *Proc Natl Acad Sci* 98:676–682.
- Rubia K, Smith AB, Brammer MJ, Taylor E (2003) Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. *NeuroImage* 20:351–358.
- Sasaki Y, Watanabe T (2017) When perceptual learning occurs. *Nat Hum Behav* 1:1–2.
- Schacht JP, Anton RF, Myrick H (2013) Functional neuroimaging studies of alcohol cue reactivity: a quantitative meta-analysis and systematic review. *Addict Biol* 18:121–133.
- Schiller B, Gianotti LR, Baumgartner T, Nash K, Koenig T, Knoch D (2016) Clocking the social mind by identifying mental processes in the IAT with electrical neuroimaging. *Proc Natl Acad Sci* 113:2786–2791.
- Slade T, Chapman C, Swift W, Keyes K, Tonks Z, Teesson M (2016) Birth cohort trends in the global epidemiology of alcohol use and alcohol-related harms in men and women: systematic review and meta-regression. *BMJ Open* 6:e011827.
- Stein M, Fey W, Koenig T, Oehy J, Moggi F (2018) Context-specific inhibition is related to craving in alcohol use disorders: a dangerous imbalance. *Alcohol Clin Exp Res* 42:69–80.
- Syrjanen E, Wiens S (2013) Gender moderates valence effects on the late positive potential to emotional distracters. *Neurosci Lett* 551:89–93.
- Tschuempferlin RM, Stein M, Batschelet HM, Moggi F, Soravia LM (2019) Learning to resist the urge: a double-blind, randomized controlled trial investigating alcohol-specific inhibition training in abstinent patients with alcohol use disorder. *Trials* 20:402.
- Volkow ND, Baler RD (2014) Addiction science: uncovering neurobiological complexity. *Neuropharmacology* 76:235–349.
- Voluse AC, Gioia CJ, Sobell LC, Dum M, Sobell MB, Simco ER (2012) Psychometric properties of the Drug Use Disorders Identification Test (DUDIT) with substance abusers in outpatient and residential treatment. *Addict Behav* 37:36–41.
- Wechsler H, Davenport A, Dowdall G, Moeykens B, Castillo S (1994) Health and behavioral consequences of binge drinking in college. A national survey of students at 140 campuses. *JAMA* 272:1672–1677.
- Wiers RW, Van De Luitgaarden J, Van Den Wildenberg E, Smulders FT (2005) Challenging implicit and explicit alcohol-related cognitions in young heavy drinkers. *Addiction* 100:806–819.
- Wiers RW, Van Woerden N, Smulders FTY, De Jong PJ (2002) Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *J Abnorm Psychol* 111:648–658.
- Williams JK, Themanon JR (2011) Neural correlates of the implicit association test: evidence for semantic and emotional processing. *Soc Cogn Affect Neurosci* 6:468–476.
- Wittchen HU, Pfister H (1997) *Instruktionsmanual zur Durchführung von DIA-X-Interviews*. Swets Test Services, Frankfurt.
- Xiao F, Zheng Z, Wang Y, Cui J, Chen Y (2015) Conflict monitoring and stimulus categorization processes involved in the prosocial attitude implicit association test: Evidence from event-related potentials. *Soc Neurosci* 10:408–417.
- Yeung N, Botvinick MM, Cohen JD (2004) The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev* 111:931–959.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Supplementary Material.** Additional information on questionnaires and procedure (SI1), behavior analyses (SI2), microstate analyses (SI3), source analyses (SI4), and analyses of gender effects in controls (SI5) is provided.